International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 8 Issue 5, Sep-Oct 2024 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Instagram Spam Detection (ISD)

Pranali V. Dhote[1], Rima Ramteke[2], Anjali R. Patel[3], Pratik Dewalker[4], Prof. Anupam Chaube[5]

[1,2,3,4]School of Science, G. H. Raisoni University, Amravati, Maharashtra, India
[5]Dean, G. H. Raisoni University, Amravati, Maharashtra, India

ABSTRACT

An Instagram spam detection project would cover the project's aim, methods, and outcomes. It would detail how the project identifies and filters out spam content on Instagram to enhance user experience and security. The abstract might mention the use of machine learning algorithms, natural language processing, and image recognition to detect and remove spam posts, comments, and accounts. It could also highlight the importance of such a project in maintaining a clean and authentic environment on the platform. The focus is on developing algorithms and tools to automatically identify and filter out spam content on the platform. This involves using machine learning techniques to analyze patterns in user behavior, text content, and images to distinguish between legitimate posts and spam. By training the system on a large dataset of known spam content, the model can learn to recognize and flag suspicious activity in real-time, helping to maintain a safe and enjoyable environment for user. Assuring a safe and fulfilling user experience on the well-known social media platform requires the crucial duty of Instagram Spam Detection (ISD). Cyberbullying, identity theft, and even financial fraud can result from unsolicited messages on Instagram. Several machine learning algorithms for Instagram spam detection have been presented by researchers as a solution to this problem. Tokenization, stop word removal, and sentiment analysis using the VADER algorithm are some methods of preparing the text data. Following preprocessing, Count Vectorizer is used to turn the data into numerical feature vectors. Based on the labeled data, three classifiers are trained and assessed: F1 score, accuracy, precision, recall, and Decision Trees/Random Forest. Taking into consideration weighted parameters that are essential in establishing an account's legitimacy, Gradient Boosting Classifier has demonstrated encouraging results in detecting phony accounts on Instagram. Instagram streaming spam detection continues to be difficult, and a strong detection method should take into account the elements of popular topics, content, URL, and user identity.

How to cite this paper: Pranali V. Dhote | Rima Ramteke | Anjali R. Patel | Pratik Dewalker | Prof. Anupam Chaube "Instagram Spam Detection (ISD)" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-8 | Issue-5, October 2024, pp.573-583, URL: www.ijtsrd.com/papers/ijtsrd69419.pdf

IJTSRD69419

KEYWORDS: *Instagram, Spam Detection (ISD), Machine Learning, Natural Language Processing (NLP), Feature Extraction, Text Preprocessing, Tokenization, Stop Word Removal, Sentiment Analysis, VADER, CountVectorizer, Naive Bayes, Support Vector Machine (SVM), Random Forest, Complementary Naive Bayes (CNB), Deep Learning*

I. INTRODUCTION

Because of the popularity of spam content on Instagram and its extensive use, spam detection on the site has become a major field of research and application. The issue of preserving a secure and pleasurable user improved in order to accommodate changing spam strategies. Enhancing the robustness of ISD solutions requires the integration of real-time detection systems and the investigation of novel aspects, like the use of emojis and contextual analysis. In conclusion, Instagram Spam Detection marks a crucial junction.

More recently, post-comment pairings and emoji features were added to improve Instagram's spam comment detection. The accuracy of spam detection was increased by employing ensemble machine learning techniques. According to the study, adding post-comment relationships and emoji features can improve spam classifier performance. ML and Deep Learning Methods Comparison.

An analysis of machine learning and deep learning methods for identifying Indonesian spam comments on Instagram was conducted.

To sum up, SVMs, Random Forest, Naive Bayes, and With Instagram's growing popularity, spam identification has become more and more important, especially when it comes to comments. The proliferation of spam—described as unnecessary, deceptive, or damaging messages—presents serious obstacles to the integrity of the platform and user experience when users interact with information. This makes the creation of efficient automated spam detection systems necessary to guarantee a secure and entertaining user experience. Several essential elements are included in the Instagram Spam Detection (ISD) architecture that is being suggested. First, information is gathered via comments left by users on a variety of posts.

II. RELATED WORK:

Detecting Spam Comments using Complementary Naive Bayes:

One method for handling imbalanced datasets for Instagram spam comment detection makes use of the Complementary Naive Bayes (CNB) algorithm. The CNB algorithm is contrasted with other methods for screening spam comments on blogs, such as K-nearest neighbor, neural networks, and support vector machines.

Assessing ML Techniques for Spam Profile Identification:

Another study evaluated the efficacy of different machine learning algorithms for Instagram spam profile detection, such as Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbor (KNN), and Multilayer Perceptron (MLP). The Random Forest algorithm fared better than the other techniques on the WEKA and RapidMiner platforms.

Improving Spam Detection with Emoji Functionality:

More recently, post-comment pairings and emoji features were added to improve Instagram's spam comment detection. The accuracy of spam detection was increased by employing ensemble machine learning techniques. According to the study, adding post-comment relationships and emoji features can improve spam classifier performance.

ML and Deep Learning Methods Comparison:

An analysis of machine learning and deep learning methods for identifying Indonesian spam comments on Instagram was conducted. Various machine learning and deep learning models were trained and assessed following the preparation of a dataset, preprocessing, and feature engineering. To sum up, SVMs, Random Forest, Naive Bayes, and SVMs, random forests, deep learning, and Naive Bayes have all demonstrated potential in identifying spam on Instagram. The performance of these spam detection systems can be further improved by adding further characteristics like emojis and post-comment context. Still, additional study is required to create reliable, in-the-moment spam filtering for social media networks.

Improving Spam Detection with Emoji Functionality:

More recently, post-comment pairings and emoji features were added to improve Instagram's spam comment detection. The accuracy of spam detection was increased by employing ensemble machine learning techniques. According to the study, adding post-comment relationships and emoji features can improve spam classifier performance.

ML and Deep Learning Methods Comparison:

An analysis of machine learning and deep learning methods for identifying Indonesian spam comments on Instagram was conducted. Various machine learning and deep learning models were trained and assessed following the preparation of a dataset, preprocessing, and feature engineering.

To sum up, SVMs, Random Forest, Naive Bayes, and Machine Learning Methodologies Various research works have utilized machine learning methods to identify spam on Instagram.

A feature-based approach for spam post detection was put forth that makes use of supervised learning strategies like K-fold cross validation. Spam and non-spam posts were categorized using well-known algorithms including Random Forest, Decision Trees, and Naive Bayes.

In particular, for unbalanced datasets, Complementary Naive Bayes (CNB) proved to be efficacious in identifying spam comments on Instagram. SVM weighting with TF-IDF was employed for comparison.

Datasets:

Introduced for spam detection research, the SPAMID-PAIR dataset comprises pairs of Instagram posts and comments from Indonesia that include emoji. Profile data for training fake profile detection models can be found in the Instagram Fake Spammer Genuine Accounts dataset from Kaggle.

III. PROPOSED WORK

Here is a suggested method that combines image analysis and comment identification, together with sample source code snippets, for Instagram spam detection in Python. You can use a variety of machine learning algorithms and packages for this purpose.

Comment Spam Detection:

Use Natural Language Processing (NLP) text analysis tools to find spam comments. This may entail using classifiers like Support Vector Machines (SVM) or

Complementary Naive Bayes (CNB) in conjunction with techniques like TF-IDF.

Python text analysis techniques combined with Natural Language Processing (NLP) and machine learning algorithms can be used to identify spam comments on Instagram. This is a suggested method.

Data Preprocessing

Open the dataset: Acquire a dataset of Instagram comments that have been classified as spam or not.

Straighten and prepare the text: Eliminate mentions, hashtags, emojis, URLs, and carry out further cleaning procedures. Text should be lowercased before being tokenized into words.

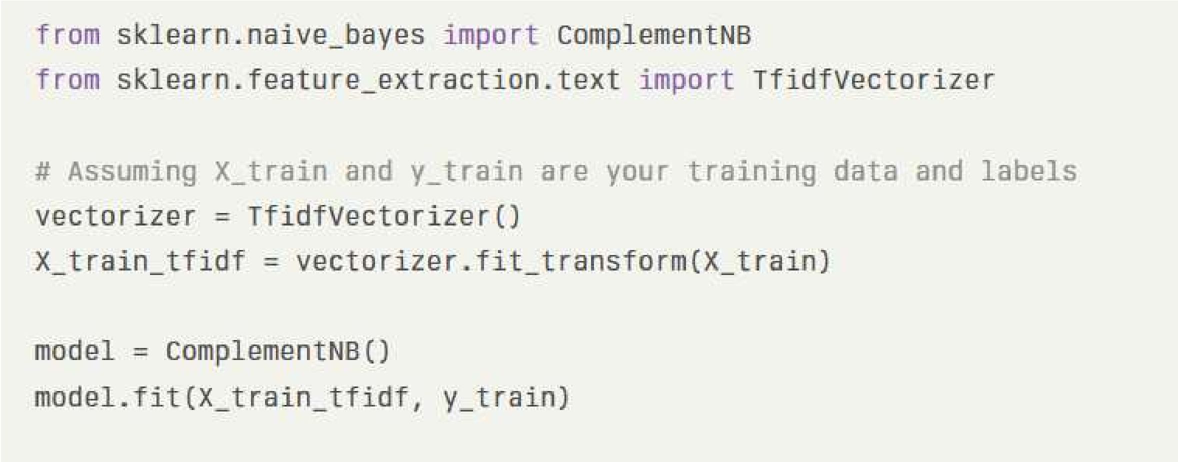
Take features out: To transform the text into numerical feature vectors, apply methods such as TF-IDF (Term Frequency-Inverse Document Frequency).

Model Training

Split the dataset: Separate the training and testing sets from the preprocessed data.

Educate a classifier: Train a spam detection model on the training set of data using machine learning algorithms such as Logistic Regression, Support Vector Machines (SVM), and Complement Naive Bayes (CNB).

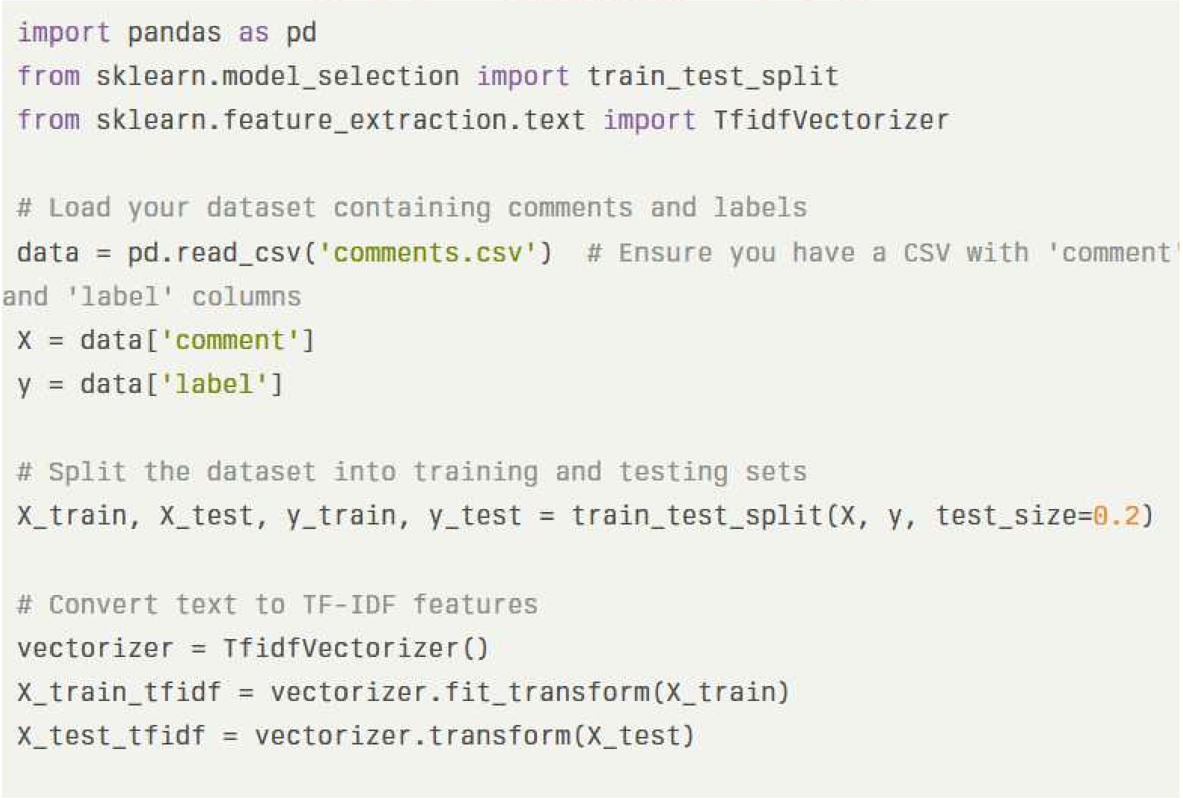
Assess the model: Utilize the testing set to assess the performance of the trained model with F1-score, accuracy, precision, and recall.



```
from sklearn.naive_bayes import ComplementNB
from sklearn.feature_extraction.text import TfidfVectorizer

# Assuming X_train and y_train are your training data and labels
vectorizer = TfidfVectorizer()
X_train_tfidf = vectorizer.fit_transform(X_train)

model = ComplementNB()
model.fit(X_train_tfidf, y_train)
```

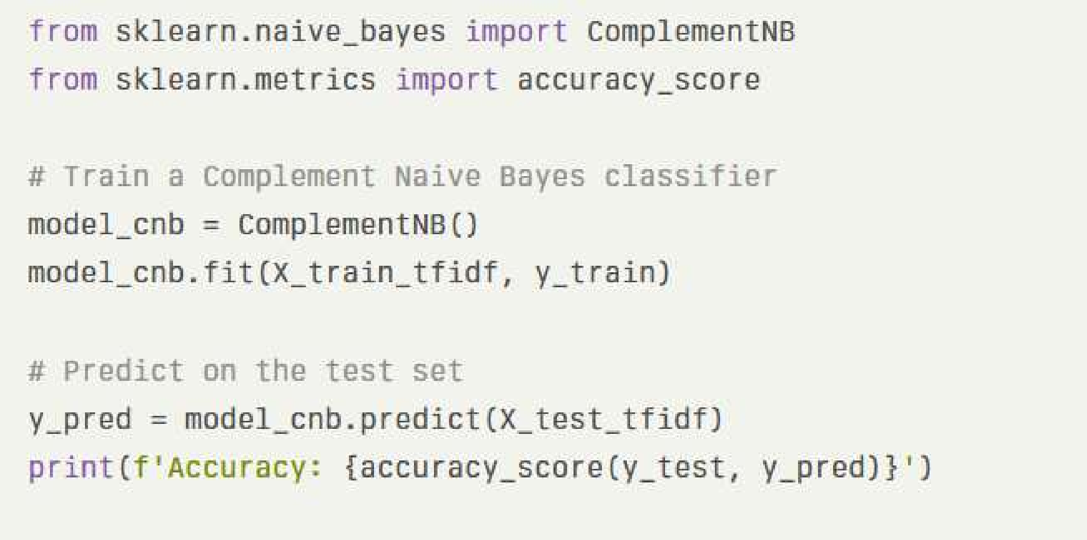


```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer

# Load your dataset containing comments and labels
data = pd.read_csv('comments.csv')  # Ensure you have a CSV with 'comment'
and 'label' columns
X = data['comment']
y = data['label']

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

# Convert text to TF-IDF features
vectorizer = TfidfVectorizer()
X_train_tfidf = vectorizer.fit_transform(X_train)
X_test_tfidf = vectorizer.transform(X_test)
```



```
from sklearn.naive_bayes import ComplementNB
from sklearn.metrics import accuracy_score

# Train a Complement Naive Bayes classifier
model_cnb = ComplementNB()
model_cnb.fit(X_train_tfidf, y_train)

# Predict on the test set
y_pred = model_cnb.predict(X_test_tfidf)
print(f'Accuracy: {accuracy_score(y_test, y_pred)}')
```

This framework offers a fundamental structure for analyzing images and comments on Instagram in order to identify spam. By adjusting the hyperparameters and utilizing more sophisticated strategies like ensemble methods or deep learning architectures customized for your dataset, you can improve the model even more.

Look into existing repositories on sites like GitHub or Kaggle for a complete implementation with extra capabilities like preserving models or effectively managing big datasets.

To prevent biased predictions, train on a balanced dataset.

To identify the top-performing model, try out several feature extraction methods and machine learning algorithms.

To increase the model's accuracy over time, add new data to it on a regular basis.

For simple Instagram integration, implement the solution as a web application or API.

IV. PROPOSED RESEARCH MODEL

Various research projects and approaches that have been offered in the literature can provide insights for conducting a system analysis for comment spam detection on Instagram. An organized summary of the essential elements needed to create a successful spam detection system can be seen below.

Instagram Comment Spam Detection

Definition of the Problem

The main objective is to recognize and remove spam comments on Instagram, which frequently contain links to harmful websites, useless messages, and advertising content. The difficulty is in differentiating between valid and spam comments, especially in datasets where the proportion of non-spam comments to spam comments is usually skewed.

Data Collection

Data can be collected from Instagram postings, focusing on user-generated comments. This dataset ought to contain a range of comments classified as non-spam or spam. For example, one study tested several categorization techniques using a dataset of 24,000 manually annotated comments from posts by Indonesian public figures.

Data Preprocessing

In order to get the data ready for analysis, preprocessing steps are essential:
Text cleaning: Take out extraneous words, punctuation, hashtags, emoticons, and URLs.

Normalization: To maintain consistency, all text should be converted to lowercase.

Divide the text into discrete words, or tokens, using tokenization.

Stop Word Removal: Remove frequently used words (such "and" and "the") that don't add anything to the text.

Feature Extraction

Feature extraction transforms the cleaned text into a format suitable for machine learning models:

Bag-of-Words (BoW): Represents text data as a matrix of token counts.

TF-IDF (Term Frequency-Inverse Document Frequency): Highlights important words in the comments based on their frequency across documents.

Word Embeddings: Techniques like fast Text or Word2Vec can capture semantic meanings of words.

Model Selection

Various machine learning algorithms can be employed to classify comments:
Complementary Naive Bayes (CNB): Particularly effective for imbalanced datasets and has shown promising results in detecting spam comments.

Support Vector Machine (SVM): A robust classifier that can be used as a benchmark against other methods.

Ensemble Methods: Combining multiple models can improve accuracy; recent studies suggest using features like emoji-text pairs to enhance detection performance.

Model Training and Evaluation

The dataset is split into training and testing subsets to evaluate model performance:
Use metrics such as accuracy, precision, recall, and F1-score to assess how well the model identifies spam comments.

K-fold cross-validation can help ensure that the model generalizes well across different subsets of data.

Implementation and Deployment

After the model is trained, it can be integrated into a service or application that keeps track of comments on Instagram in real-time. To do this, use API calls to retrieve fresh comments from Instagram posts. Preprocess incoming comments by applying the same pipeline that was used for training. Then, use the trained model to categorize each comment and take the necessary action (like hiding or deleting spam).

Continuous Improvement

To maintain effectiveness against evolving spam tactics:

Regularly update the dataset with new examples of spam comments.

Retrain the model periodically to adapt to changes in user behavior and spamming techniques.

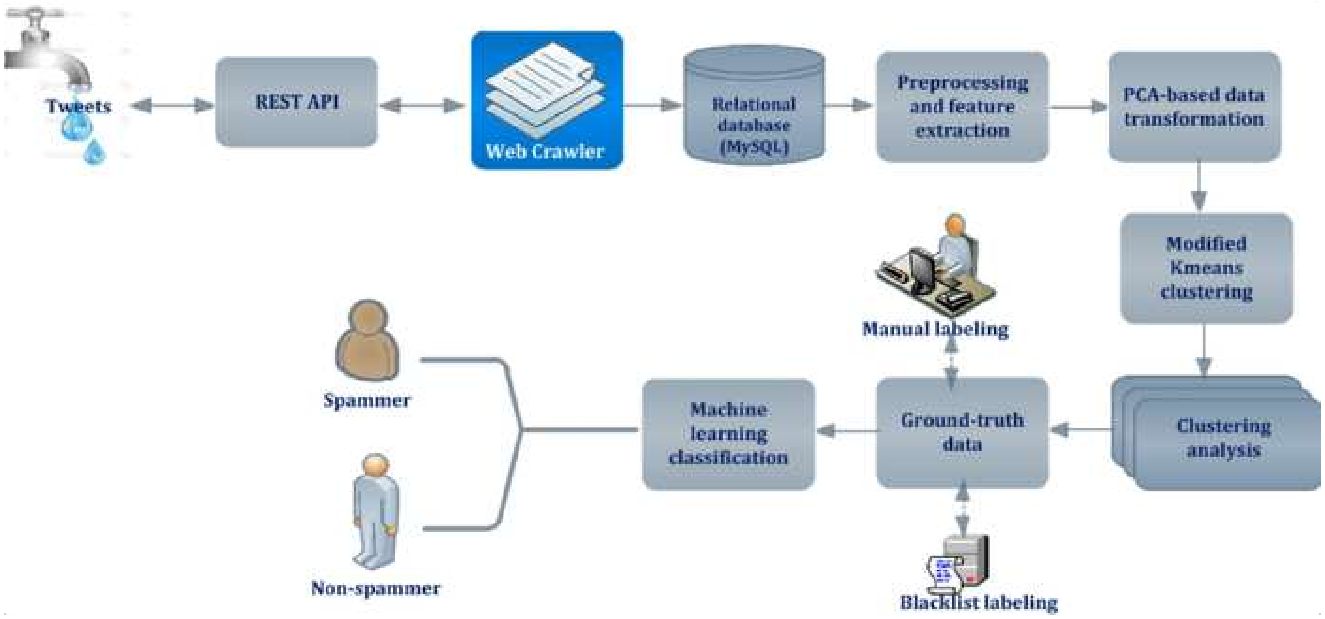

Tweets
REST API
Web Crawler
Relational database (MySQL)
Preprocessing and feature extraction
PCA-based data transformation
Modified Kmeans clustering
Manual labeling
Spammer
Machine learning classification
Ground-truth data
Clustering analysis
Non-spammer
Blacklist labeling

Fig 1 : Comment Spam Detection

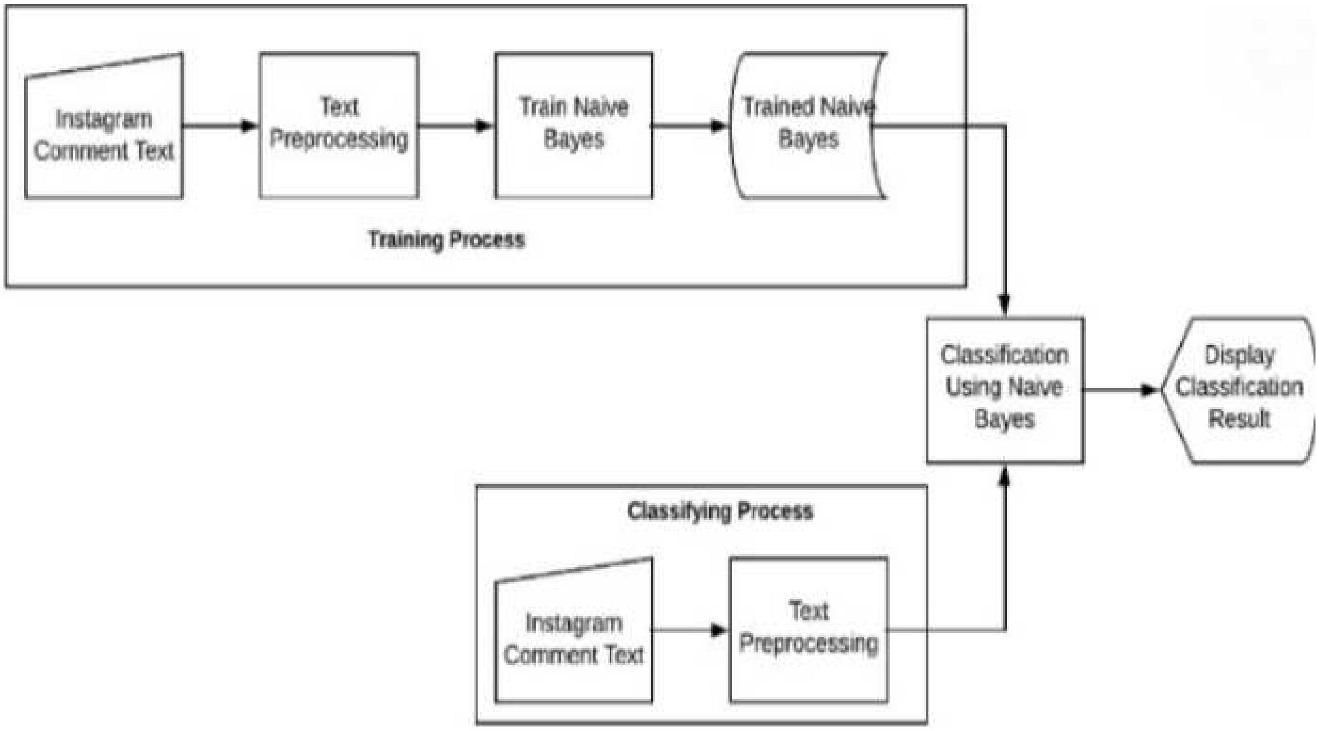


Fig 2 : Comment Spam Detection on Instagram

V. PERFORMANCE EVALUATION

Implementing and Expanding

To make deployment and scaling simple, package the services using containerization (e.g., Docker).

For scalability and high availability, deploy the services on a cloud platform (such as AWS, Google Cloud, or Azure).

To make operations simpler, use managed services for machine learning, database management, and data processing.

To accommodate more traffic, provide horizontal scaling for the Real-Time Spam Detection Service.

Enhance response speeds for real-time spam classification by utilizing caching methods.

Privacy and Security

Use HTTPS and authentication techniques to establish secure communication between services.

By adhering to best practices for data storage and access management, you can protect data privacy.

Keep an eye out for security flaws in the system and update it frequently.

Observation and Recordkeeping

Enable thorough recording for every service to facilitate troubleshooting and debugging.

Logging and metrics should be gathered, stored, and visualized using a logging and monitoring solution (such as Elasticsearch, Kibana, Prometheus, or Grafana).

Configure alerts for important occasions and benchmarks in performance.

Based on knowledge from numerous research and approaches, a number of tactics can be used to enhance an Instagram spam detection model's performance. This is a methodical technique.

You may create a reliable and effective comment spam detection system for Instagram by adapting best practices for scalability, security, and maintainability into your system architecture.

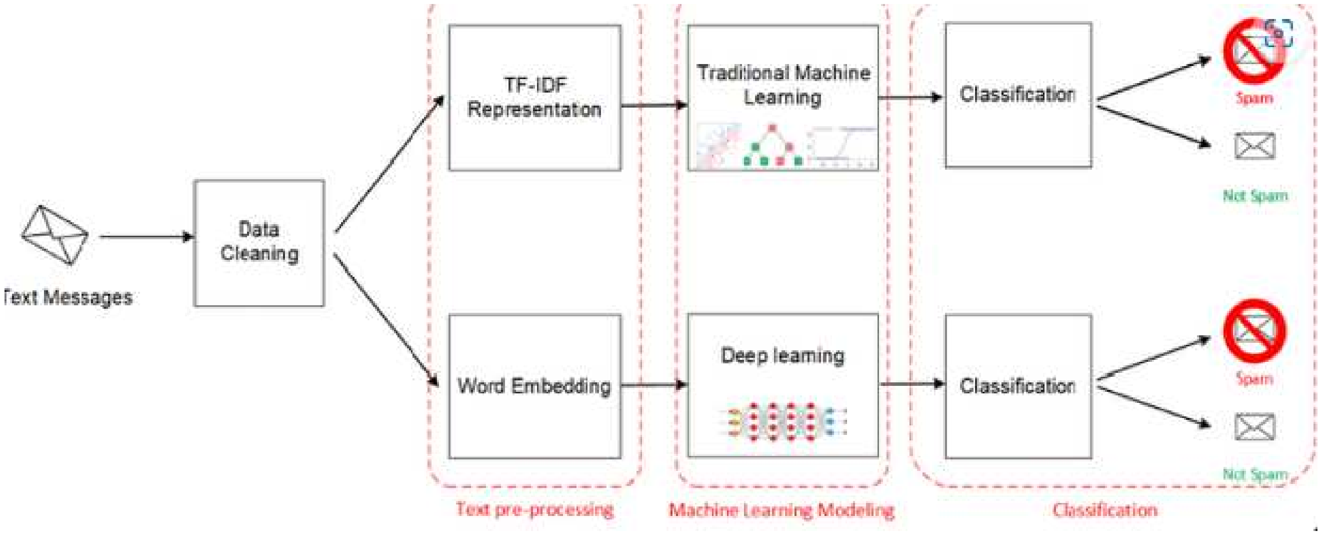
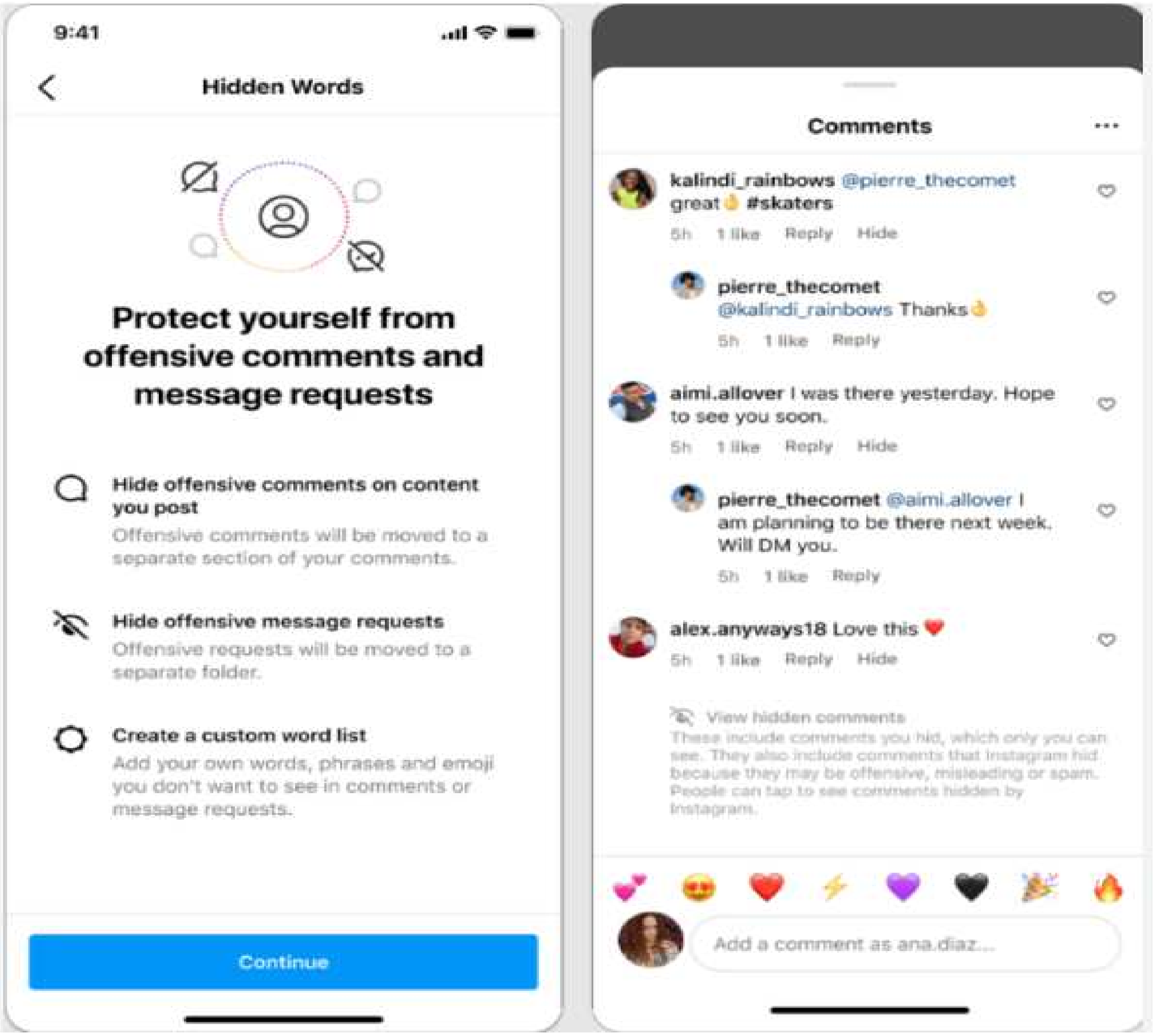


Fig 3 : Architecture of the spam detection model



9:41

Hidden Words

Protect yourself from offensive comments and message requests

Hide offensive comments on content you post
Offensive comments will be moved to a separate section of your comments.

Hide offensive message requests
Offensive requests will be moved to a separate folder.

Create a custom word list
Add your own words, phrases and emoji you don't want to see in comments or message requests.

Continue

Comments

kalindi_rainbows @pierre_thecomet great👌 #skaters
5h 1 like Reply Hide

pierre_thecomet @kalindi_rainbows Thanks👌
5h 1 like Reply

aimi.allover I was there yesterday. Hope to see you soon.
5h 1 like Reply Hide

pierre_thecomet @aimi.allover I am planning to be there next week. Will DM you.
5h 1 like Reply

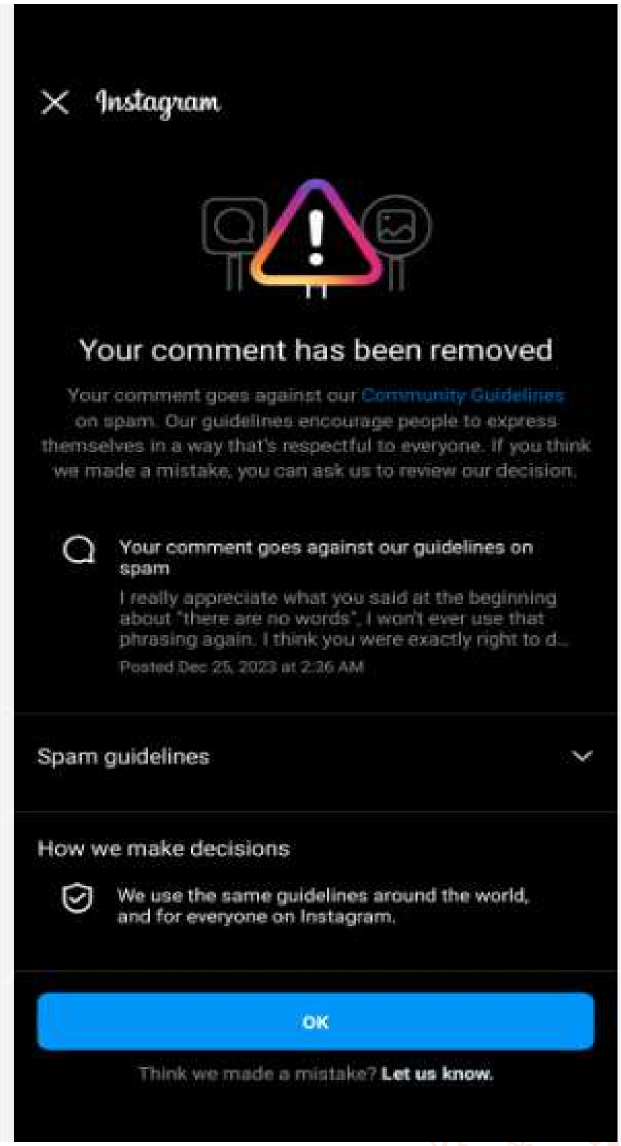
alex.anyways18 Love this ❤️
5h 1 like Reply Hide

View hidden comments
These include comments you hid, which only you can see. They also include comments that Instagram hid because they may be offensive, misleading or spam. People can tap to see comments hidden by Instagram.

💕 😍 ❤️ ⚡ 💜 🖤 🎉 🔥

Add a comment as ana.diaz...

Fig 4 : Protection from unwanted comments



Instagram

Your comment has been removed

Your comment goes against our Community Guidelines on spam. Our guidelines encourage people to express themselves in a way that's respectful to everyone. If you think we made a mistake, you can ask us to review our decision.

Your comment goes against our guidelines on spam

I really appreciate what you said at the beginning about "there are no words", I won't ever use that phrasing again. I think you were exactly right to d...

Posted Dec 25, 2023 at 2:36 AM

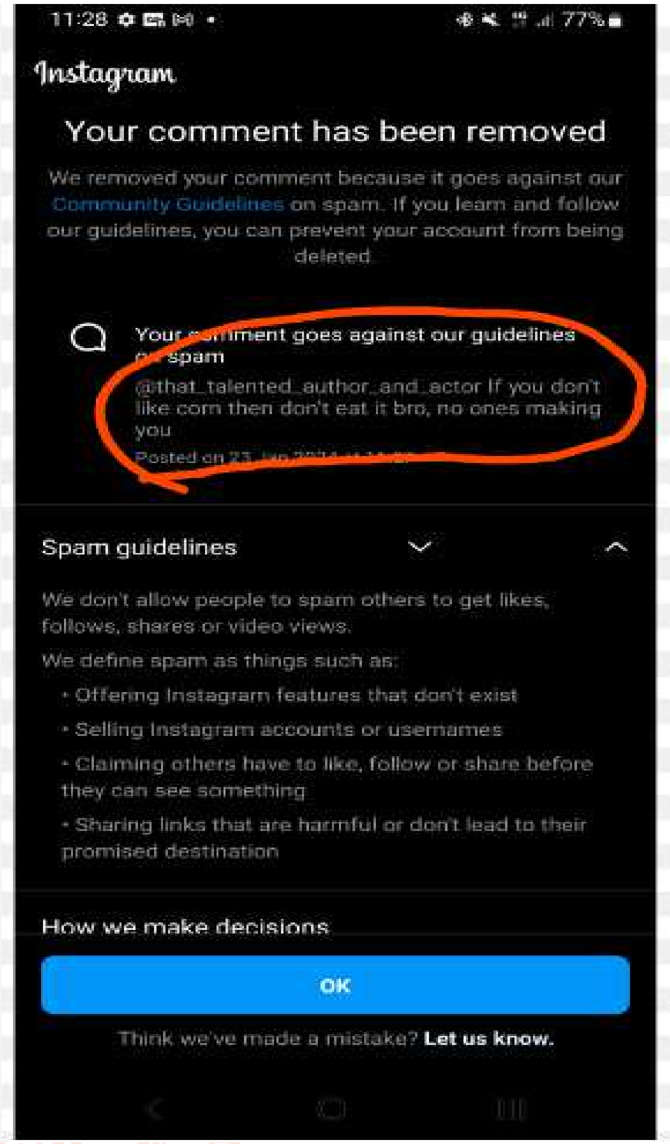
Spam guidelines

How we make decisions

We use the same guidelines around the world, and for everyone on Instagram.

OK

Think we made a mistake? **Let us know.**



11:28 77%

Instagram

Your comment has been removed

We removed your comment because it goes against our Community Guidelines on spam. If you learn and follow our guidelines, you can prevent your account from being deleted.

Your comment goes against our guidelines on spam

@that_talented_author_and_actor If you don't like corn then don't eat it bro, no ones making you

Posted on 23 [illegible]

Spam guidelines

We don't allow people to spam others to get likes, follows, shares or video views.

We define spam as things such as:

- Offering Instagram features that don't exist
- Selling Instagram accounts or usernames
- Claiming others have to like, follow or share before they can see something
- Sharing links that are harmful or don't lead to their promised destination

How we make decisions

OK

Think we've made a mistake? **Let us know.**

Fig 5 :Your comments has been removed **Fig 6 :Banned for this comments**

VI. RESULT ANALYSIS

Feature extraction approaches, performance evaluation measures, and machine learning techniques are combined in the analysis of Instagram spam detection, especially for comments. Numerous strategies and their efficacy in spotting spam comments on the site have been emphasized by recent studies.

Measures of Performance

Several performance measures are used to assess the efficacy of spam detection algorithms:

Accuracy: The model's overall correctness.

Accuracy and Memory: Recall evaluates the model's capacity to find all pertinent cases, whereas precision estimates the percentage of true positives among anticipated positives.

F1 Points: By striking a compromise between recall and precision, this metric offers a single score for assessing model performance.

Studies employing sophisticated models, such as those based on BERT architectures or ensemble techniques, have, for example, reported F1 values more than 0.93.

Graphical Representation

Although the search results do not include specific graphs, common graphical representations of spam detection research could look somewhat like this:

To see true positives, false positives, true negatives, and false negatives, use confusion matrices.

ROC curves: Showing how specificity (false positive rate) and sensitivity (true positive rate) are traded off at various thresholds.

Bar charts: Displaying F1 or accuracy ratings for different feature sets or algorithms.

These visual aids aid in comprehending the variations in model performance as well as the effects of different characteristics on detecting skills.

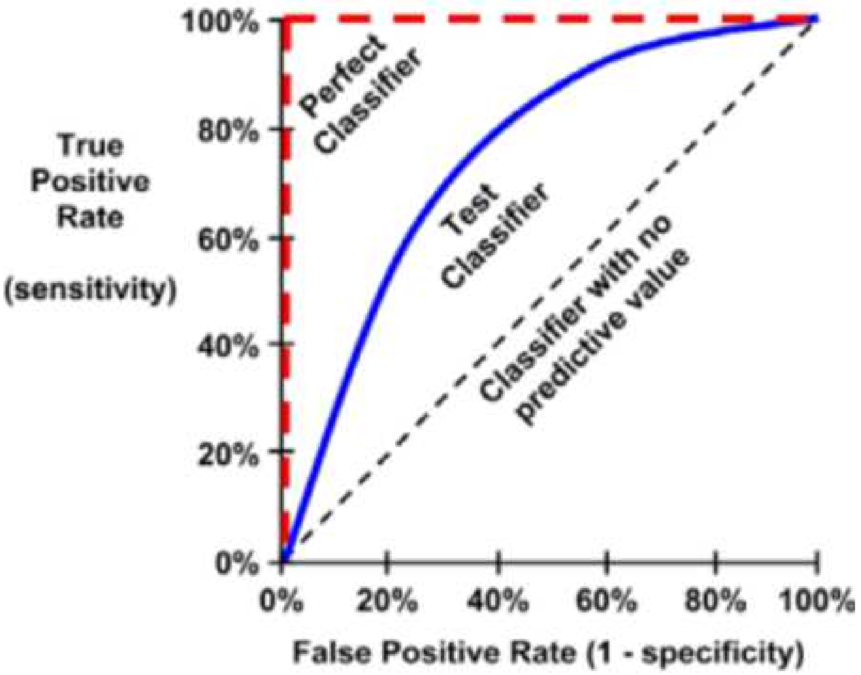


Fig: ROC (Receiver – Operating Characterstic) curves

The necessity of combining cutting-edge machine learning algorithms with efficient feature extraction approaches is shown by the continuing research on Instagram comment spam detection. Accurately recognizing spam comments has been demonstrated through the use of models such as Complementary Naïve Bayes and ensemble techniques. In order to make these models more resilient to changing spam strategies, future research will probably continue to improve them by adding larger datasets and more complex features.

VII. CONCLUSION

In conclusion, preserving a great user experience and protecting the integrity of the platform depend on efficient comment spam detection on Instagram. Instagram can effectively identify and remove spammy content, such as repetitive comments, pointless links, and criminal activity, by utilizing machine learning algorithms and natural language processing. When user reporting mechanisms are combined with real-time detection technologies, the result is a dynamic solution that can adjust to changing spam tactics. To prevent penalizing legitimate users, it's crucial to achieve a balance between lowering false positives and spam identification. Furthermore, privacy protection and ethical issues must be taken into account when designing and implementing these systems. In general, a strong and flexible spam detection system is essential to guaranteeing a more secure and interesting Instagram environment. All things considered, using machine learning to identify spam comments on Instagram offers a workable solution to a common issue with social media management. Platforms can minimize the negative effects of spam on community interactions while greatly enhancing user experiences by continuously improving detection algorithms and incorporating user feedback.

One major issue that impacts user experience and engagement on Instagram is the detection of bogus comments. Numerous research have shown that spam comments may be successfully identified and filtered out by using machine learning approaches, including Complementary Naive Bayes (CNB) and other sophisticated algorithms.

VIII. REFERENCE

[1] Databooks, "Ini Media Sosial Paling Populer Sepanjang April 2020," Databooks, 2020. https://databoks.katadata.co.id/datapublish/2020/05/25/ini-media-sosial-paling-populer-sepanjang-april-2020 (accessed Nov. 04, 2020).

[2] S. Aiyar and N. P. Shetty, "N-Gram Assisted Youtube Spam Comment Detection," Procedia

Computer Science., vol. 132, pp. 174–182, 2018, doi: 10.1016/j.procs.2018.05.181.

[3] A. R. Chrismanto, A. K. Sari, and Y. Suyanto, "CRITICAL EVALUATION ON SPAM CONTENT DETECTION IN SOCIAL MEDIA," Journal of Theoretical and Applied Information Technology (JATIT), vol. 100, no. 8, pp. 2642–2667, 2022, [Online]. Available: http://www.jatit.org/volumes/Vol100No8/29Vol100No8.pdf

[4] A. Chrismanto and Y. Lukito, "Klasifikasi Komentar Spam Pada Instagram Berbahasa Indonesia Menggunakan K-NN," in Seminar Nasional Teknologi Informasi Kesehatan vv (SNATIK), 2017, pp. 298–306.

[5] F. Prabowo and A. Purwarianti, "Instagram online shop's comment classification using statistical approach," in Proceedings - 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2017, 2018, pp. 282–287. doi:10.1109/ICITISEE.2017.8285512.

[6] A. Chrismanto and Y. Lukito, "Deteksi Komentar Spam Bahasa Indonesia Pada Instagram Menggunakan Naive Bayes," Jurnal Ultima, vol. 9, no. 1, pp. 50–58, 2017, doi:10.31937/ti.v9i1.564.

[7] W. Zhang and H.-M. Sun, "Instagram Spam Detection," in 2017 IEEE 22nd Pacific Rim International Symposium on Dependable Computing (PRDC), Jan. 2017, pp. 227–228. doi: 10.1109/PRDC.2017.43.

[8] B. Priyoko and A. Yaqin, "Implementation of naive bayes algorithm for spam comments classification on Instagram," in 2019 International Conference on Information and Communications Technology, ICOIACT 2019, 2019, pp. 508–513. doi:10.1109/ICOIACT46704.2019.8938575.

[9] N. A. Haqimi, N. Rokhman, and S. Priyanta, "Detection Of Spam Comments On Instagram Using Complementary Naïve Bayes," IJCCS (Indonesian Journal of Computing and Cybernetics Systems, vol. 13, no. 3, p. 263, Jul. 2019, doi: 10.22146/ijccs.47046.

[10] A. Chrismanto and Y. Lukito, "Identifikasi Komentar Spam Pada Instagram," Lontar Komputer: Jurnal Ilmiah Teknologi Informasi, vol. 8, no. 3, p. 219, 2017, doi:10.24843/lkjiti.2017.v08.i03.p08.

[11] A. Chrismanto, Y. Lukito, and A. Susilo, "Implementasi Distance Weighted K-Nearest Neighbor Untuk Klasifikasi Spam dan Non-Spam Pada Komentar Instagram," Jurnal Edukasi dan Penelitan Informatika, vol. 6, no. 2, p. 236, 2020, doi: 10.26418/jp.v6i2.39996.

[12] A. Chrismanto, W. Raharjo, and Y. Lukito, "Design and Development of REST-Based Instagram Spam Detector for Indonesian Language," Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, iSemantic 2018, iSemantic 2018, pp. 345–350, Sep. 2018, doi:10.1109/ISEMANTIC.2018.8549725.

[13] A. R. Chrismanto, W. Sudiarto, and Y. Lukito, "Integration of REST-Based Web Service and Browser Extension for Instagram Spam Detection," International Journal of Advanced Computer Science and Applications, vol. 9, no. 12, 2018, doi:10.14569/IJACSA.2018.091253.

[14] C. Zhang, C. Liu, X. Zhang, and G. Almpanidis, "An up-to-date comparison of state-of-the-art classification algorithms," Expert Systems with Applications., vol. 82, pp. 128–150, 2017, doi:10.1016/j.eswa.2017.04.003.

[15] M. P. Nugraha, A. Nurhadiyatna, and D. M. S. Arsa, "Offline Signature Identification Using Deep Learning and Euclidean Distance," Lontar Komputer : Jurnal Ilmiah Teknologi Informasi, vol. 12, no. 2, pp. 102–111, Aug. 2021, doi: 10.24843/LKJITI.2021.V12.I02.P04.

[16] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "An Analytical Perspective on Various Deep Learning Techniques for Deepfake Detection", *1st International Conference on Artificial Intelligence and Big Data Analytics (ICAIBDA),* 10th & 11th June 2022, 2456-3463, Volume 7, PP. 25-30, https://doi.org/10.46335/IJIES.2022.7.8.5

[17] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "Revealing and Classification of Deepfakes Videos Images using a Customize Convolution Neural Network Model", *International Conference on Machine Learning and Data Engineering (ICMLDE)*, 7th & 8th September 2022, 2636-2652, Volume 218, PP. 2636-2652, https://doi.org/10.1016/j.procs.2023.01.237

[18] Usha Kosarkar, Gopal Sakarkar (2023), "Unmasking Deep Fakes: Advancements,

Challenges, and Ethical Considerations", *4th International Conference on Electrical and Electronics Engineering (ICEEE)*,19th & 20th August 2023, 978-981-99-8661-3, Volume 1115, PP. 249-262, https://doi.org/10.1007/978-981-99-8661-3_19

[19] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2021), "Deepfakes, a threat to society", *International Journal of Scientific Research in Science and Technology (IJSRST)*, 13th October 2021, 2395-602X, Volume 9, Issue 6, PP. 1132-1140, https://ijsrst.com/IJSRST219682

[20] Usha Kosarkar, Prachi Sasankar(2021), " A study for Face Recognition using techniques PCA and KNN", Journal of Computer Engineering (IOSR-JCE), 2278-0661,PP 2-5,

[21] Usha Kosarkar, Gopal Sakarkar (2024), "Design an efficient VARMA LSTM GRU model for identification of deep-fake images via dynamic window-based spatio-temporal analysis", Journal of Multimedia Tools and Applications, 1380-7501, https://doi.org/10.1007/s11042-024-19220-w

[22] Usha Kosarkar, Dipali Bhende, "Employing Artificial Intelligence Techniques in Mental Health Diagnostic Expert System", International Journal of Computer Engineering (IOSR-JCE),2278-0661, PP-40-45, https://www.iosrjournals.org/iosr-jce/papers/conf.15013/Volume%202/9.%2040-45.pdf?id=7557